When Google Translate is better than Some Human Colleagues, those People are no longer Colleagues

Läubli, Samuel ; Orrego-Carmona, David

Abstract: We analyse posts on social media (Facebook, LinkedIn, and Twitter) as a means to understand how translators feel about machine translation (MT). A quantitative analysis of more than 13,000 tweets shows that negative perceptions outweigh positive ones by a ratio of 3:1 overall, and 5:1 in tweets relating MT to human translation. Our study indicates a disconnect between translation and research communities, and we outline three suggestions to bridge this gap: (i) identifying and reporting patterns rather than isolated errors, (ii) participating in evaluation campaigns, and (iii) engaging in cross-disciplinary discourse. Rather than pointing out each other's deficiencies, we call for computer scientists, translation scholars, and professional translators to advance translation technology by acting in concert.

Originally published at:
Läubli, Samuel; Orrego-Carmona, David (2017). When Google Translate is better than Some Human Colleagues, those People are no longer Colleagues. In: Translating and the Computer, London, 16 November 2017 - 17 November 2017. AsLing, 59-69.

When Google Translate is better than Some Human Colleagues, those People are no longer Colleagues

Samuel Läubli[1] and **David Orrego-Carmona**[2,3]

[1]Institute of Computational Linguistics, University of Zurich, Switzerland
[2]School of Languages and Social Sciences, Aston University, United Kingdom
[3]Department of Linguistics and Language Practice, University of the Free State, South Africa

Abstract

We analyse posts on social media (Facebook, LinkedIn, and Twitter) as a means to understand how translators feel about machine translation (MT). A quantitative analysis of more than 13,000 tweets shows that negative perceptions outweigh positive ones by a ratio of 3:1 overall, and 5:1 in tweets relating MT to human translation. Our study indicates a disconnect between translation and research communities, and we outline three suggestions to bridge this gap: (i) identifying and reporting patterns rather than isolated errors, (ii) participating in evaluation campaigns, and (iii) engaging in cross-disciplinary discourse. Rather than pointing out each other's deficiencies, we call for computer scientists, translation scholars, and professional translators to advance translation technology by acting in concert.

1 Introduction

Mistranslations can be hilarious. In fact, social media have become ideal outlets to share pictures of clumsy food menus and mislabelled street signs, as well as screenshots of translation errors produced by machine translation (MT) engines such as Google Translate. People share them, laugh at them, and criticise them openly. Professional translators also participate in the debates about such translations. On a regular basis, translators on LinkedIn, Facebook and Twitter engage in discussions about mistranslations and how they show that MT is not comparable to human translation (see Figure 1). Translators use these spaces to voice their frustration with MT and the implications it has on their profession.

Social media groups dedicated to translation and translators count their members in the thousands. Communities of practice have emerged thanks to these spaces; however, researchers have barely looked at them in order to better understand translators and their opinions. Considering the lack of attention to translators' activities on social media and curious about how these could be used to understand the translators' perceptions of MT, we decided to conduct a study into how translators' interactions on social media could help improving translation technology.

Perceptions of MT among translators have been explored using questionnaires and interviews. We conjectured that eliciting their opinions from online

Figure 1: Meme posted on a translators' group on Facebook, mocking the use of Google Translate.

Proceedings of the 39th Conference Translating and the Computer, pages 59–69,
London, UK, November 16-17, 2017. ©2017 AsLing

interactions would provide us with data to
understand their attitudes towards MT, and propose ways in which their efforts and knowledge could support the improvement of the technology.

We used qualitative and quantitative methods to analyse how translators' feel about MT. We first present the results of our initial qualitative exploration of groups on Facebook and LinkedIn. The posts on these platforms gave us the impression that sentiment towards MT in translation groups is predominantly negative. Aiming at quantifying this initial impression and providing empirical grounding, we then employed automatic sentiment analysis on a larger data set. We classified a collection of 13,150 tweets about MT using human annotation on a subsample of 150 tweets, and automatic annotation for the entire collection.[1]

Both our qualitative and quantitative analyses show that negative perceptions in social media outnumber positives. To the best of our knowledge, these results provide the first empirical view of how MT is portrayed on social media. Based on our findings, we make a call for improving collaboration among professional translators and researchers, and propose possible avenues to move towards that goal.

2 Background

Most of the literature on the perception of MT among translators, some of which we review in this section, relies on data obtained through formal questionnaires and interviews. This paper is motivated by our impression that translators might be more open and direct when expressing opinions on social media, as well as the fact that there is a lot more data than could be collected through direct interrogation.

Already in 1993, Meijer found that the MT was seen as a threat among translators, and negative opinions seem to persist (see Guerberof Arenas, 2013; Gaspari et al., 2015; Cadwell et al., 2017). Despite significant technological advancements in recent years, translators are 'still strongly resistant to adopting MT as an aid, and have a considerable number of concerns about the impact it might have on their long-term work practices and skills' (Cadwell et al., 2017). As a response to these concerns, in the last two years, the International Federation of Translators (FIT) has published three position papers on MT, crowdsourcing, and the future for professional translators. In their paper on MT, they state that 'MT is unlikely to completely replace human translators in the foreseeable future. Leaving aside the area where MT is a feasible option, there will continue to be plenty of work for them. Professional translators, who have the appropriate skills and qualifications, will still be needed for demanding, high-quality products' (Heard, 2017). Given that the Federation represents the interests of professional translators, their paper can be seen as an indicator of the relevance to understand how translators feel about MT.

In spite of the seemingly significant importance for the community,[2] the use of social media among professional translators has been barely studied. Desjardins (2016) addresses the aspects of professionals using social media but primarily as a strategy to increase their visibility, not as a way of interacting among themselves. The research that is available in the field of translation and social media has mainly explored the work of non-professionals and their translations (e. g., Dombek, 2014; O'Hagan, 2017; Jiménez-Crespo, 2017). Although not on social media, the online presence of translators and their attitudes in other outlets have previously been

[1] All data and annotations are released under the CC BY-SA 4.0 license, available at `https://github.com/laeubli/MTweet`.

[2] Social media groups dedicated to translation and translators count their members in the thousands, and since 2013, `proz.com` has been running the Community Choice Awards to recognise, among others, translation and interpreting professionals and companies who are active and influential on the internet.

explored: McDonough Dolmaya (2011) analysed translators' blog entries to understand the attitudes and practices of professional translators, while Flanagan (2016) used blogs to study their opinions towards crowdsourcing. Along the same lines, researchers have also asked professional translators about their attitude towards MT (Meijer, 1993), their opinion about post-editing of MT output (Guerberof Arenas, 2013), and their reasons to use or not use MT (Cadwell et al., 2017).

The research on the translators' opinions about MT is still limited and, to the best of our knowledge, no study has analysed interactions on social media as a way of understanding the translators' attitude towards MT. Interacting on social media requires less time and effort than maintaining a website or writing a blog post so we assume a larger number of translators would be involved in different types of exchanges in social media platforms.

3 Qualitative Analysis

Our aim to fill this gap started with a preliminary analysis of translators' posts and comments on Facebook and LinkedIn. We hand-picked 137 examples related directly or indirectly to MT by browsing through public and invitation-only[3] groups: Professional Translators and Interpreters (ProZ.com),[4] Translation Quality,[5] The League of Extraordinary Translators,[6] Things Translators Never Say,[7] and Translation and Interpreting Group.[8] It is important to point out here that this part of the study does not claim to be comprehensive; it serves to illustrate the situation that unfolds in these groups rather than provide generalisable results.

In relation to the assumption that MT can be a threat to professional translators, one of the recurrent topics in these groups is quality. Translators engage in discussions about the mistranslations produced by MT engines as a way of reinforcing the need for human translators. Photos and screenshots of translation errors are systematically posted to the groups. Translators criticise them and comment on the shortcomings of MT (see Figure 2b). Some use the examples to respond with sarcasm to the possibility that translators might be replaced by machines in the near future: 'Oh yes, I'm very worried about being replaced by a machine when it can't tell the difference between an extraordinary announcement and a declaration of emergency...'. In their discussion, Google Translate is normally the main culprit, probably because of its accessibility and the considerable number of languages in which it operates. There are direct references to Google Translate in 66 of the 137 posts we collected for this qualitative analysis.

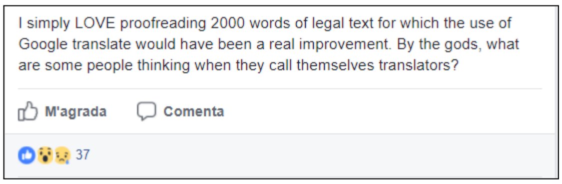
Translators also question the improvements announced by the companies that develop MT. In response to an article comparing the quality of neural MT to phrase-based MT and human translation, one translator indicates her doubts about the results commenting 'I wonder how the quality was measured if neural came so close to human.' In this case, MT as such is not the issue that is put on the spot but the concept of quality that is used to assess MT output. Also, as pointed out by other researchers (see Doherty and Kenny, 2014; Cadwell et al., 2017), translators feel they are not considered part of the development of MT: 'Yes, AI people who know nothing about our job, we totally agree that you will figure out how to replace us with machines in the next ten years. Sure you will.'

In some cases, translators even use MT as an indicator of poor quality when judging other translators and their translations. Figure 2d shows a comment by a translator who uses Google Translate's output quality as a point of comparison to argue that the translation she

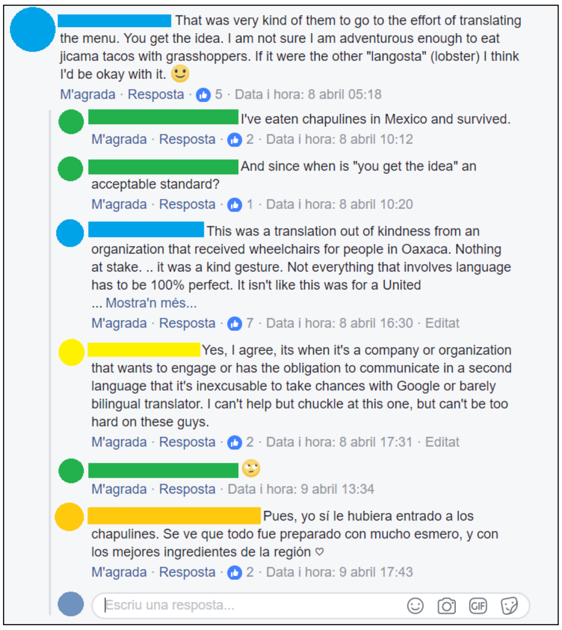
[3] Access is usually granted within minutes.
[4] `https://www.linkedin.com/groups/138763`
[5] `https://www.linkedin.com/groups/3877235`
[6] `https://www.facebook.com/groups/extraordinarytranslators`
[7] `https://www.facebook.com/groups/thingstranslatorsneversay`
[8] `https://www.facebook.com/groups/Interpreting.and.Translation`

(a) Translation from German into English made by Google Translate used as an example of how confusing MT outputs can be. In the comments, the translators discuss Google Translate's poor attempt at rendering the different meanings of the terms in German. In the output in English, all the German terms are translated as 'economics', resulting in a meaningless repetition of the same term.

(b) A translator posted a link to a list of examples of mistranslations generated by Google Translate App's function that can translate text in images. The translator sarcastically comments on the fact that the poor quality of the translations makes it unlikely that human translators will be replaced by computers in the near future.



(d) A translator complaints about the quality of the translation she is revising. As a way of signalling the poor quality of the translation made by her colleague, she claims it would have been better to proofread a machine-translated text.



(c) Translators engage in a discussion about whether or not MT can be acceptable in specific circumstances. Some of them argue MT can be useful for small business without the resources to pay for a professional translation, while others stress the fact that accepting MT as a valid option means lowering the standards of the profession.

Figure 2: Examples of translator interactions in Facebook groups.

is proofreading is of low quality. Translators also recognise that some of the mistakes present in the translations that are posted in the group are such poor examples of translations that 'not even Google translate [sic] is that bad.'

However, not all the posts and comments on social media discredit MT straight away. Figure 2a presents an image that was shared in one of the groups showing Google Translate's translation of a German text into English. Interestingly, the translators who commented on this post were genuinely curious about the veracity of the output. Some of them took the time to retype the text into Google Translate and check whether the translations into English or their own target languages made any sense.

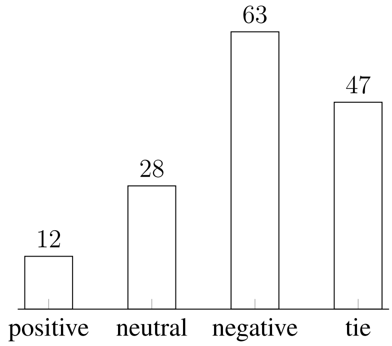
Comments in the groups often also point at the use of MT as an aid for translating as an indicator of poor quality or a poorly skilled professional. One of the commentators states that 'Machine translation, like Google Translate, can give you a false sense of competence', suggesting that non professionals could get the impression they can translate thanks to the support of MT. Another translator comments on the fact that the fear of MT is, in a way, an indicator of the competence of the translators. She says that '[m]achines will only replace those translators who translate like machines.' These opinions do not represent isolated cases. In another thread when discussing the issues that MT could bring to the profession, a translator states that 'When Google Translate is better than some human colleagues, those people are no longer colleagues.' Using Google Translate or the risk of being replaced by a machine seem then to be related to a translator's lack of professionalism or skills.

One of the highlights of MT is affordability: automating the process of translating makes it possible for people to access translations, even when they do not have the resources to pay for them. The discussion depicted in Figure 2c serves as an example of this argument among translators. Some of the translators recognise there are situations in which having access to an automatic translation is better than having no translation at all, while others would not consider it possible to accept a translation that only allows users to 'get the idea'. For some of the translators, it seems, accepting MT as a valid option would constitute lowering the standards of the profession.

Discussions in the groups commonly go back to the assumption that human translators approach translation as a creative task, while MT only looks at translation as the word-for-word replacement of a string of text. Not all the discussions centre on the negative aspects of MT. Some translators point out that MT, and Google Translate in particular, are good for certain language combinations or specific fields, and can actually support the work of skilled professionals. A translator summarises these two points when he states that 'Translation and interpreting are very demanding professions where talented human linguists will continue to make the difference in value and quality. Nevertheless, it is hard to deny the benefits of applied language technology – CAT for translators and VRI for interpreters to name but a few – to support linguists and language service providers in their joint mission to meet customer requirements in a very rapidly changing market of demanding end users and organizations who pay the bill for these language services.'

4 Quantitative Analysis

The initial exploration of how MT is discussed on social media reinforced our impression that perceptions are predominantly negative among professional translators. We conducted a larger study in order to ground this impression empirically. In this stage, we focused on Twitter data as large numbers of posts are difficult to obtain from Facebook and LinkedIn (see Section 4.1). Our goal was to quantify the extent of positive, neutral, and negative tweets on MT, for which we employed independent human judges (Section 4.2) and an automatic sentiment classifier



(a) Assignments per class

		Annotator B		
	$n = 150$	positive	neutral	negative
Annotator A	positive	**12**	11	6
	neutral	11	**28**	2
	negative	8	9	**63**

(b) Confusion matrix

Figure 3: Human Sentiment Analysis

(Section 4.3).

4.1 Data Collection

We collected tweets from `twitter.com` using a purpose-built, open source web crawler.[9] We only kept tweets which (i) contain the terms 'machine translation' and/or 'machine translated', (ii) are written in English, according to Twitter's language identification, and (iii) were created between 1 January 2015 and 31 July 2017. This method is not exhaustive in that authors may refer to machine translation by use of synonyms or without mentioning it explicitly. However, the data is representative of what a user would find searching for the terms mentioned in (i) above through Twitter's search interface. Our filtered collection contains 13,150 tweets.

4.2 Human Sentiment Analysis

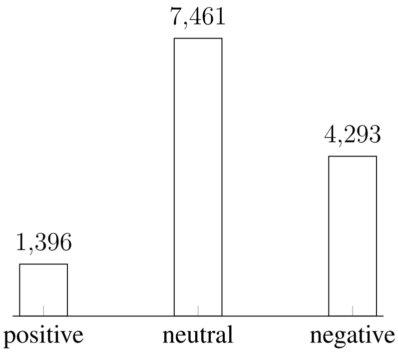
We sampled 150 tweets from this collection for human sentiment analysis. The selection was random, except that we required each tweet to contain at least one of the following terms: 'human', 'professional', 'translator'. As discussed in Section 5.2, we used this heuristic to focus on discussions comparing MT and human translation in this part of the study.

The sampled tweets formed the basis for an annotation job on a web-based crowdsourcing platform.[10] Annotators were asked to read each tweet, click all links found in the text for additional context, and then determine if the tweet is positive, neutral, or negative with regards to machine translation. Tweets were presented in random order. We included ten control items as a means to filter out random contributions: each annotator saw ten tweets twice, and we expected them to be consistent with their own judgements.

Human annotators were recruited through convenience sampling, the restriction being that they have never been involved with translation, translation studies, or computational linguistics. Five annotators completed the entire job, from which we excluded three due to low inter-annotator agreement: they failed to reproduce their own judgements on three or more out of ten control items.

Human annotation results are summarised in Figure 3. Inter-annotator agreement is 68.7 % (Cohen's $\kappa = 0.495$). The two remaining annotators independently assigned the same label to 103 out of 150 tweets: 12 positive, 28 neutral, and 63 negative (Figure 3a). Two different labels were assigned to 47 tweets, with most disagreement between positive and neutral (Figure 3b).

[9] `https://github.com/jonbakerfish/TweetScraper`
[10] `https://www.crowdflower.com`



(a) Assignments per class

		Automatic		
	$n = 103$	positive	neutral	negative
Human	positive	**2**	4	6
	neutral	3	**21**	4
	negative	1	17	**45**

(b) Confusion matrix

Figure 4: Automatic Sentiment Analysis

4.3 Automatic Sentiment Analysis

To annotate our entire collection of tweets (see Section 4.1), we leveraged Baziotis et al.'s (2017) automatic sentiment classifier.[11] Their system scored first in the SemEval 2017 shared task on classifying the overall sentiment of a tweet (task 4, subtask A), i. e., deciding whether it expresses positive, neutral or negative sentiment (see Rosenthal et al., 2017). It uses a deep LSTM network (Hochreiter and Schmidhuber, 1997) with attention (Rocktäschel et al., 2015) and is completely data-driven: rather than relying on linguistic information and hand-crafted rules, the system learns to classify from large collections of manually annotated tweets.

We trained the system on the original SemEval data, meaning it is not specifically geared to tweets on machine translation. Using the 103 tweets that both of our annotators labelled with the same class as a reference (see Section 4.2), it classifies 68 tweets correctly. This corresponds to an overall classification accuracy of 66.0 %. In terms of average recall per class – the primary evaluation metric used in SemEval – its performance in our domain (66.0 %) is similar to the performance achieved in the shared task with varied topics (68.1 %; see Rosenthal et al., 2017). However, precision (33.3 %) and recall (16.7 %) are low for positive tweets. The system performs better with neutral (precision: 50.0 %, recall: 75.0 %) and negative tweets (precision: 81.8 %, recall: 71.4 %), with a tendency to classify negative tweets as neutral (Figure 4b).

Overall, the classifier labels 1,396 tweets as positive (10.6 %), 7,461 as neutral (56.7 %), and 4,293 as negative (32.6 %). Note that in contrast to the subset used for human annotation (see Section 4.2), this includes tweets not comprising the terms 'human', 'professional', or 'translator'.

5 Findings and Discussion

Our study provides evidence that MT is often portrayed negatively among translators on social media outlets. The suspicions about a negative attitude towards MT that stemmed from our qualitative analysis of Facebook and LinkedIn posts (Section 3) were supported by the results of the sentiment analysis carried out on Twitter data (Section 4).

5.1 Recurrent Topics

Our exploration of Facebook and LinkedIn data (see Section 3) sheds light on recurrent MT-related topics in social media. Firstly, we observed frequent reiteration of how professional

[11] source code available at `https://github.com/cbaziotis/datastories-semeval2017-task4`

translators are and will still be needed as MT improves, as shown by the example is provided in Figure 2b.

Secondly, translators doubt if MT improves at all, for example, by calling into question the methodology and/or veracity of evaluation campaigns. Referring to a study on productivity with a web-based translation workbench, a Facebook user says 'if only your productivity estimate was correct! If I actually could do 2k words/hour while watching esports [sic], I'd actually take on all those bottom feeders and still make good bank!'

Thirdly, many posts merely criticise MT for bad quality. Translators spend considerable time and effort on discussing MT errors, but we were surprised to find little variance in the discussions and errors reported. In one instance, a translator even made up a meaningless sentence in Japanese and mocks Google Translate for producing meaningless output, conceivably because of its sexual connotation (see Figure 5).

5.2 Sentiment Towards MT

In analysing 150 tweets relating MT to human translation, two independent judges found negative tweets to be most common, outnumbering positive and neutral ones by a ratio of 5:1 and 2:1, respectively. However, sentiment classification in tweets is not trivial: human judgements do not overlap in a third of all cases, resulting in 47 ties. As shown in Table 1c, there are even tweets classified as positive by one annotator and negative by the other. Even if other studies on human sentiment analysis in tweets report similar inter-annotator agreement (e. g., Cieliebak et al., 2017), a negotiation phase following the independent procedure of annotation could have resolved some of the disagreement.

Moreover, sampling tweets based on the presence of keywords – 'human', 'professional', or 'translator' – is somewhat arbitrary. Still, we found this heuristic useful to get a sense of how Twitter users contrast MT with human translation (see the examples in Table 1).

Without this restriction, neutral tweets on MT are most common in our collection. News and product announcements, such as 'Facebook posts its fast and accurate ConvNet models for machine translation on GitHub', often fall into this category. But even so, there are three times more negative than positive tweets in the 13,150 examples we collected, hinting at the predominance of negative perceptions about MT in general.

The caveat here is that sentiment was determined by means of an automatic classifier. The classifier did not have access to contents such as websites and images linked in tweets, which human annotators were explicitly asked to consider when making their judgement. It also was not geared to tweets on MT specifically; while the system we leveraged would have allowed for topic-based classification (see Baziotis et al., 2017), we lacked appropriate amounts of training data. Despite these limitations, the system reproduced human judgements with an accuracy of 66.0 % overall. This corresponds to state-of-the-art results (see Rosenthal et al., 2017), and is similar to the degree of disagreement between human annotators (see above). This is good enough to get a sense of the class distribution in our data, even if the classifier does make mistakes (e. g., Table 1b). A clear advantage is speed: the 13,150 tweets are labelled in seconds. Eliciting human annotations would have taken a lot longer and would have been expensive at this scale.

5.3 A Case for Collaboration

Even after its emergence as a profession in the last century, translation still struggles with recognition and undervaluation (see Tyulenev, 2015; Flanagan, 2016). Apart from the general situation of the profession, translators also feel, and indeed in many cases are, left out in the processes towards the development of translation technologies (see Doherty and Kenny,

	Text	Automatic	Human	Human$_A$	Human$_B$
(a)	Six reasons why machine translation can never replace good human translation: https://t.co/JzLYbXO6yJ #xl8 #t9n	negative	negative	negative	negative
(b)	When you solely rely on machine translation... via @inspirobot #wedoitthehumanway #htt https://t.co/UpfnVd4k8W	neutral	negative	negative	negative
(c)	High-quality machine translation is threatening to make translators 'the coffee-bean pickers of the future' https://t.co/n8fGvIHBao	negative	tie	positive	negative
(d)	Difference between professional translation and machine translation by @ChrisDurbanFR #xl8 #ITIConf17 https://t.co/gFhgRrLtJq	neutral	neutral	neutral	neutral
(e)	Pretty incredible. For a few languages, machine translation is near equal to human translation. https://t.co/GsCejE0cUW	positive	positive	positive	positive

Table 1: Example tweets. Sentiment was assessed by two human annotators as well as an automatic sentiment classifier.

2014; Cadwell et al., 2017). They resist technology because they feel they need to protect their profession and resort to the defence of quality as the main argument for their cause. This behaviour is neither a new strategy, nor something restrictive of professional translators. As Pym (2011) puts it: 'Resistance to technological change is usually a defense of old accrued power, dressed in the guise of quality.' However, it is unlikely that the technological development will stop. Together with its quality, the use of MT has increased significantly in the last decades. Research has also provided evidence of increased productivity through post-editing of MT, and companies are moving more and more towards a context in which this practice is the norm rather than the exception (e. g., Green et al., 2013; Koponen, 2016). While it can be assumed that translation will continue being an activity with human involvement, it will (continue to) involve various degrees of automation as translation technologies evolve. We believe that translators should be actively engaged in these developments, and that their actions on social media could help inform and support research on translation technology. In the following section, we propose a set of recommendations aimed at fostering collaboration and promoting common goals among researchers and professional translators.

6 Recommendations

As hilarious as MT errors can be, laughing about them does neither improve translators' lives nor the technology. The study we present in this paper fills a gap in the exploration and quantification of translators' perceptions as it brings social media into the picture. Our findings imply that translators and researchers have different understandings of the functioning and purposes of MT, but at the same time show that translators are aware of the types of issues that are problematic for it. Considering our findings, we believe that professional translators could and should have more influence on future developments in MT and translation technology in general, and propose three initial recommendations to bridge this gap:

Figure 5: Translation of a meaningless text in Japanese into English by Google Translate. The translator posts the screenshot because Google Translate's output has a sexual connotation; he uses this argument as proof that machines will never replace human translators.

(i) *Identify and report patterns rather than isolated errors.* State-of-the-art MT systems are not based on rules, so complaining about specific words that are mistranslated does not help much. Reporting error patterns, in contrast, may help gearing systems to new objectives, such as ensuring noun phrase agreement or controlling for negation particles (see also Sennrich, 2017). Professional translators have the knowledge and expertise to identify these patterns and, given the appropriate tools, pattern reporting could easily be integrated into their regular workflows.

(ii) *Participate in evaluation campaigns.* Our study has shown that criticising findings of MT quality or productivity as being unrealistic is a recurrent theme on social media. The MT research community puts a lot of effort into evaluation campaigns. At the annual Workshop on Machine Translation (WMT), for example, research teams from all over the world compete for the best translation system in several language pairs. Human judgements are the primary evaluation metric, but rather than from translators, they stem from 'researchers who contributed evaluations proportional to the number of [MT systems] they entered' into the competition (Bojar et al., 2017) – although blindly, the MT researchers evaluate their own systems. We believe that the involvement of professional translators in scientific evaluation campaigns would not only improve the quality and credibility of their outcome, but also improve the translators' understanding of – and impact on – the methodologies used to evaluate MT systems.

(iii) *Engage in cross-disciplinary discourse.* We need to talk to each other. The issues presented above show that professional translators and researchers hold different positions, not due to a lack of information or skills, but rather poor communication. Translators make good points about the shortcomings of MT, but they comment on MT issues primarily among themselves. The outcomes of academic research, on the other hand, are poorly communicated to translators and seemingly give them the impression that MT has little to offer to improve their conditions. For translators to have an active role in the development of the technologies they (have to) use, it is necessary for both sides, professional translators and researchers, to meet halfway and cooperate. Common rather than separate spaces on social media may be a starting point.

References

Baziotis, Christos, Nikos Pelekis, and Christos Doulkeridis. 2017. Datastories at SemEval-2017 task 4: Deep LSTM with attention for message-level and topic-based sentiment analysis. In *Proceedings of the 11th International Workshop on Semantic Evaluation*. Vancouver, Canada, pages 747–754.

Bojar, Ondřej, Rajen Chatterjee, Christian Federmann, Yvette Graham, Barry Haddow, Shujian Huang, Matthias Huck, Philipp Koehn, Qun Liu, Varvara Logacheva, Christof Monz, Matteo Negri, Matt Post, Raphael Rubino, Lucia Specia, and Marco Turchi. 2017. Findings of the 2017 conference on machine translation (WMT17). In *Proceedings of the Second Conference on Machine Translation, Volume 2: Shared Task Papers*. Copenhagen, Denmark, pages 169–214.

Cadwell, Patrick, Sharon O'Brien, and Carlos S. C. Teixeira. 2017. Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators. To appear in *Perspectives*, available online at `http://www.tandfonline.com/doi/full/10.1080/0907676X.2017.1337210`.

Cieliebak, Mark, Jan Milan Deriu, Dominic Egger, and Fatih Uzdilli. 2017. A Twitter corpus and benchmark resources for German sentiment analysis. In *Proceedings of the 5th International Workshop on Natural Language Processing for Social Media*. Valencia, Spain, pages 45–51.

Desjardins, Renée. 2016. *Translation and Social Media: In Theory, in Training and in Professional Practice*. Palgrave Macmillan.

Doherty, Stephen and Dorothy Kenny. 2014. The design and evaluation of a statistical machine translation syllabus for translation students. *The Interpreter and Translator Trainer* 8(2):295–315.

Dombek, Magdalena. 2014. *A study into the motivations of internet users contributing to translation crowdsourcing: the case of Polish Facebook user-translators*. Ph.D. thesis, Dublin City University.

Flanagan, Marian. 2016. Cause for concern? Attitudes towards translation crowdsourcing in professional translators' blogs. *The Journal of Specialised Translation* 25:149–173.

Gaspari, Federico, Hala Almaghout, and Stephen Doherty. 2015. A survey of machine translation competences: insights for translation technology educators and practitioners. *Perspectives* 23(3):333–358.

Green, Spence, Jeffrey Heer, and Christopher D. Manning. 2013. The efficacy of human post-editing for language translation. In *Proceedings of the 2013 Conference on Human Factors in Computing Systems (CHI)*. Paris, France.

Guerberof Arenas, Ana. 2013. What do professional translators think about post-editing? *The Journal of Specialised Translation* 19:75–95.

Heard, Reiner. 2017. FIT position paper on machine translation. *Babel* 63(1):130–135.

Hochreiter, Sepp and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9(8):1735–1780.

Jiménez-Crespo, Miguel A. 2017. *Crowdsourcing and online collaborative translations: Expanding the limits of Translation Studies*, volume 131 of *Benjamins Translation Library*. John Benjamins.

Koponen, Maarit. 2016. Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation* 25:131–148.

McDonough Dolmaya, Julie. 2011. A window into the profession: What translation blogs have to offer translation studies. *The Translator* 17(1):77–104.

Meijer, Siety. 1993. Attitudes towards machine translation. *Language International* 5(6):11–13.

O'Hagan, Minako. 2017. Deconstructing translation crowdsourcing with the case of a Facebook initiative. In Dorothy Kenny, editor, *Human Issues in Translation Technology*, Routledge, pages 25–44.

Pym, Anthony. 2011. What technology does to translating. *Translation & Interpreting* 3(1):1.

Rocktäschel, Tim, Edward Grefenstette, Karl Moritz Hermann, Tomáš Kočiský, and Phil Blunsom. 2015. Reasoning about entailment with neural attention. *arXiv preprint* arXiv:1509.06664.

Rosenthal, Sara, Noura Farra, and Preslav Nakov. 2017. SemEval-2017 task 4: sentiment analysis in Twitter. In *Proceedings of the 11th International Workshop on Semantic Evaluation*. Vancouver, Canada, pages 502–518.

Sennrich, Rico. 2017. How grammatical is character-level neural machine translation? Assessing MT quality with contrastive translation pairs. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. Valencia, Spain, pages 376–382.

Tyulenev, Sergey. 2015. Towards theorising translation as an occupation. *Asia Pacific Translation and Intercultural Studies* 2(1):15–29.